October, 1981

LIDS-P-1154

TOWARD A THEORY OF NONLINEAR STOCHASTIC REALIZATION

Anders Lindquist*, Sanjoy Mitter†, and Giorgio Picci§

Feedback and Synthesis of Linear and Nonlinear Systems
Proceedings of the Workshop in
Bielefeld, West-Germany, June 22-26, 1981, and
Rome, Italy, June 29-July 3, 1981.
Edited by A. Isidori and D. Hinrichsen
Springer-Verlag Lecture Notes

---

* Department of Mathematics, University of Kentucky, Lexington, Kentucky 40506. This research was supported partially by the National Science Foundation under grant ECS-7903731 and partially by the Air Force Office of Scientific Research under grant AFOSR 78-3519.

† Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139. This research was supported by the Air Force Office of Scientific Research under grant AFOSR 77-3281D.

§ LADSEB-CNR, Corso Stati Uniti 4, 35100 Padova, Italy.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **OCT 1981**	2. REPORT TYPE	3. DATES COVERED **00-10-1981 to 00-10-1981**

4. TITLE AND SUBTITLE **Toward a Theory of Nonlinear Stochastic Realization**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Massachusetts Institute of Technology,Laboratory for Information and Decision Systems,77 Massachusetts Avenue,Cambridge,MA,02139-4307**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**16**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

TOWARD A THEORY OF NONLINEAR STOCHASTIC REALIZATION

Anders Lindquist*, Sanjoy Mitter†, and Giorgio Picci§

1. INTRODUCTION

The following is a central problem in stochastic systems theory: Given a stationary stochastic process $\{y(t);t\in\mathbb{R}\}$, find a (possibly infinite-dimensional) vector Markov process $\{x(t);t\in\mathbb{R}\}$, called the *state process*, and a function f so that $y(t) = f(x(t))$ for all $t\in\mathbb{R}$. Moreover, find a stochastic differential equation driven by a Wiener process and having the state process x as its unique solution. The problem of characterizing the family of all such representations is known as the *stochastic realization problem*.

There is by now a rather comprehensive theory of stochastic realization for the case that $\{y(t);t\in\mathbb{R}\}$ is *Gaussian* [1-3], in which case the representations can be taken to be *linear*, i.e. both f and the stochastic differential equation are linear. This linear theory can be applied to *non-Gaussian* processes also, but then we need to give up the requirement that x is Markov and that it is generated by a Wiener process, replacing these concepts by "wide sense Markov" [4] and "orthogonal increment process" respectively. If we are not willing to do so, a *nonlinear* stochastic realization theory is needed. That is the topic of this paper.

In this paper we shall apply Wiener's theory of homogeneous chaos [5,6] to the nonlinear stochastic realization problem. For simplicity and ease of notation we shall assume that the process y is scalar, although the machinary which we develop is sufficient to accommodate also the vector case. Other assumptions, such as y admitting an innovation representation, are however crucial to our approach. (In this respect, it might be more appropriate to consider a process y with stationary increments, and indeed with minor modifications we could have done so.) In the extension of this work we see the possibility of making contact with *nonlinear filtering* [7,8] and that is partially a motivation for this work.

2. PROBLEM FORMULATION

Let $\{y(t);t\in\mathbb{R}\}$ be a non-Gaussian stationary stochastic process which is mean-square continuous, purely nondeterministic, and centered, and let $\mathcal{Y}$ be the sigma-field

---

* Department of Mathematics, University of Kentucky, Lexington, Kentucky 40506. This research was supported partially by the National Science Foundation under grant ECS-7903731 and partially by the Air Force Office of Scientific Research under grant AFOSR 78-3519.

† Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139. This research was supported by the Air Force Office of Scientific Research under grant AFOSR 77-3281D.

§ LADSEB-CNR, Corso Stati Uniti 4, 35100 Padova, Italy.

generated by y. Then define H to be the Hilbert space of all centered $\mathcal{Y}$-measurable random variables, having inner product $\langle\xi,\eta\rangle = E\{\xi\eta\}$. Since y is stationary there is a strongly continuous group of unitary operators $\{U_t; t\in\mathbb{R}\}$ on H, called the *shift*, such that $y(t+s) = U_t y(s)$ for all t and s [9]. Let $\mathcal{Y}_t^-$ and $\mathcal{Y}_t^+$ be the sigma-fields generated by $\{y(s); s\le t\}$ and $\{y(s); s\ge t\}$ respectively.

Next assume that y has an innovation process $\{\nu(t); t\in\mathbb{R}\}$, by which we shall here mean a Wiener process such that $\sigma\{\nu(\tau)-\nu(\sigma); \tau,\sigma\le t\} = \mathcal{Y}_t^-$. (Here $\sigma\{\cdot\}$ denotes the sigma-field generated by the random variables inside the curly brackets.) Then, by symmetry, it also has a backward innovation process $\{\bar{\nu}(t); t\in\mathbb{R}\}$, i.e. another Wiener process such that $\sigma\{\bar{\nu}(\tau)-\bar{\nu}(\sigma); \tau,\sigma\ge t\} = \mathcal{Y}_t^+$. Now, since $\mathcal{Y} = \sigma\{\nu(t); t\in\mathbb{R}\} = \sigma\{\bar{\nu}(t); t\in\mathbb{R}\}$ we can apply Wiener's homogeneous chaos theory [5,6]. Let H_1 denote the Gaussian space [5] generated by $\{\nu(t); t\in\mathbb{R}\}$ or, which is equivalent, by $\{\bar{\nu}(t); t\in\mathbb{R}\}$. Since y is mean-square continuous, H_1 is a separable space, and therefore H_1 has a countable orthonormal basis $\{\xi_i\}_{i=0}^{\infty}$. Now, let P_n be the (closed) linear subspace of all polynomials in $\{\xi_i\}_{i=0}^{\infty}$ of degree not exceeding n. Next define $H_n = P_n \ominus P_{n-1}$, i.e. the orthogonal complement of P_{n-1} in P_n. Then it can be shown [5,6] that

$$H = H_1 \oplus H_2 \oplus H_3 \oplus \dots \tag{1}$$

where $\oplus$ denotes orthogonal direct sum. The space H_n is called the n^{th} *homogeneous chaos* of H. Since $y(0)\in H$, there is an orthogonal decomposition

$$y(0) = y_1(0) + y_2(0) + y_3(0) + \dots \tag{2}$$

where $y_n(0)\in H_n$. It is easy to see that each chaos H_n is invariant under the shift U_t, and consequently, for any $t\in\mathbb{R}$, we have a decomposition such as (2) for $y(t)$ in terms of $y_n(t) := U_t y_n(0)$, $n=1,2,3\dots$

In order to obtain a state space description we introduce a *past space* H^- and a *future space* H^+ as follows. Let $H^-(H^+)$ be the subspace of all centered $\mathcal{Y}^-$-measurable ($\mathcal{Y}^+$-measurable) random variables. Then, defining H_1^- and H_1^+ to be the Gaussian spaces generated by $\{\nu(\tau)-\nu(\sigma); \tau,\sigma\le 0\}$ and $\{\bar{\nu}(\tau)-\bar{\nu}(\sigma); \tau,\sigma\ge 0\}$ respectively, we obtain the chaos expansions

$$\begin{cases} H^- = H_1^- \oplus H_2^- \oplus H_3^- \oplus \dots & \text{(3a)} \\ H^+ = H_1^+ \oplus H_2^+ \oplus H_3^+ \oplus \dots & \text{(3b)} \end{cases}$$

where clearly $H_n^-\subset H_n$ and $H_n^+\subset H_n$ for all n. Note that $H_1^-\cap H_1^+ \ne \emptyset$ and $H_1^-\vee H_1^+ = H_1$, but $H^-\vee H^+ \ne H$.

Now, if y were a Gaussian process, y would have a component only in the first chaos, i.e. $y = y_1$, and consequently state spaces for y could be constructed along the lines of [1,2] by finding the minimal Markovian (H_1^-,H_1^+)-splitting subspaces in H_1. [We recall that, for two subspaces A and B, X is an (A,B)-*splitting subspace* if $\langle E^X\alpha, E^X\beta\rangle = \langle\alpha,\beta\rangle$ for all $\alpha\in A$ and $\beta\in B$, where E^X denotes orthogonal projection on the subspace X.] However, for a non-Gaussian process y, there will be some nontrivial component y_n, $n>1$, and consequently the state space construction will have to involve at least those higher chaoses in which y has a component. To this end define the index set

$N := \{n \mid y_n(0) \neq 0\} \cup \{1\}$. For reasons which will soon be evident, we shall have to always include the first chaos in our analysis. (In particular, see Section 7.)

Hence we call X a *state space for* y if

$$X = \bigoplus_{n \in N} X_n \ , \tag{4}$$

where $X_n \subset H_n$ is an (H_n^-, H_n^+)-splitting subspace, and X is Markovian in the sense that, if $X := \sigma\{X\}$, $X^- := \sigma\{\bigvee_{t \le 0} U_t X\}$ and $X^+ := \sigma\{\bigvee_{t \ge 0} U_t X\}$, X^- and X^+ are conditionally independent given X; we shall write this $X^- \perp\!\!\!\perp X^+ \mid X$. We say that X is *minimal* if there is no other state space X' for which $X' := \sigma\{X'\}$ is properly contained in X.

The problem at hand is now to construct all minimal state spaces for y and to obtain a dynamical representation (realization) for each of them.

3. THE STRUCTURE OF H

According to Itô's Theorem [10]

$$H_n = \{I_n(f;\nu) \mid f \in \hat{L}_2(\mathbb{R}^n)\} \tag{5}$$

where I_n is the multiple Wiener integral

$$I_n(f;\nu) = \int_{-\infty}^{\infty} \int_{-\infty}^{t_1} \cdots \int_{-\infty}^{t_{n-1}} f(t_1, t_2, \ldots, t_n)\, d\nu(t_1) \ldots d\nu(t_n) \tag{6}$$

and $\hat{L}_2(\mathbb{R}^n)$ are the symmetric functions in $L_2(\mathbb{R}^n)$. Although the region of integration is such that (5) remains the same if $\hat{L}_2(\mathbb{R}^n)$ is exchanged for $L_2(\mathbb{R}^n)$, we prefer the former since we have a one-one correspondence between elements in H_n and $\hat{L}_2(\mathbb{R}^n)$. In fact, we can establish an isometric isomorphism between these spaces [5,6,10]. Now, for $i=1,2,\ldots,n$, let $\eta_i \in H_1$ be arbitrary. Then there exist unique functions $f_i \in L_2(\mathbb{R})$ such that $\eta_i = \int_{-\infty}^{\infty} f_i(t)\, d\nu(t)$. Next define

$$\eta_1 * \eta_2 * \cdots * \eta_n = n!\ I_n(f;\nu) \tag{7}$$

where

$$f(t_1, t_2, \ldots, t_n) = \frac{1}{n!} \sum_{\pi \in G} f_{\pi_1}(t_1) f_{\pi_2}(t_2) \ldots f_{\pi_n}(t_n) \ , \tag{8}$$

G being the symmetric group of permutations of n letters. Since finite linear combinations of functions of type (8) are dense in $\hat{L}_2(\mathbb{R}^n)$, Itô's Theorem implies that

$$H_n = \overline{\text{sp}}\{\eta_1 * \eta_2 * \cdots * \eta_n \mid \eta_1, \eta_2, \ldots, \eta_n \in H_1\} \tag{9a}$$

where $\overline{\text{sp}}$ denotes closed linear hull. We shall write this as

$$H_n = H_1 * H_1 * \cdots * H_1 \ . \tag{9b}$$

By Itô's formula [11; p.38]

$$\eta_1 * \eta_2 * \cdots * \eta_n = (\eta_2 * \eta_3 * \cdots * \eta_n) \cdot \eta_1$$
$$- \sum_k (\eta_2 * \cdots * \eta_{k-1} * \eta_{k+1} * \cdots * \eta_n) \cdot \langle \eta_1, \eta_k \rangle \tag{10}$$

which can be solved recursively. For example,

$$\eta_1 * \eta_2 = \eta_1\eta_2 - \langle\eta_1,\eta_2\rangle$$

$$\eta_1 * \eta_2 * \eta_3 = \eta_1\eta_2\eta_3 - \eta_1 \cdot \langle\eta_2,\eta_3\rangle - \eta_2 \cdot \langle\eta_1,\eta_3\rangle - \eta_3 \cdot \langle\eta_1,\eta_2\rangle \,.$$

The $*$-operation is obviously commutative. In particular,

$$\underbrace{\eta * \eta * \ldots * \eta}_{n \text{ times}} = h_n(\eta, \langle\eta,\eta\rangle^{\frac{1}{2}}) \tag{11}$$

(in the sequel we shall write this η^{n*}) where

$$h_n(x,\sigma) = \frac{(-\sigma)^n}{n!} \exp\left(\frac{x^2}{2\sigma}\right) \frac{\partial^n}{\partial x^n} \exp\left(-\frac{x^2}{2\sigma}\right) , \tag{12}$$

$n=0,1,2,\ldots$, are the *Hermite polynomials* (cf [11; p.37]). Analogously to (9) we have

$$\begin{cases} H_n^- = H_1^- * H_1^- * \ldots * H_1^- & \text{(13a)} \\ H_n^+ = H_1^+ * H_1^+ * \ldots * H_1^+ & \text{(13b)} \end{cases}$$

Let $H_1^{n\circledast} = H_1 \circledast H_1 \circledast \ldots \circledast H_1$ denote the symmetric tensor-product Hilbert space of H_1 by itself taken n times. Then for arbitrary $\xi_i, \eta_i \in H_1$, $i=1,2,\ldots,n$, with $\langle\cdot,\cdot\rangle_{n\circledast}$ the inner product in $H_1^{n\circledast}$, we have

$$\langle\xi_1 \circledast \xi_2 \circledast \ldots \circledast \xi_n, \eta_1 \circledast \eta_2 \circledast \ldots \circledast \eta_n\rangle_{n\circledast} = \frac{1}{n!} \sum_{\pi\in G} \langle\xi_{\pi_1},\eta_1\rangle\langle\xi_{\pi_2},\eta_2\rangle \ldots \langle\xi_{\pi_n},\eta_n\rangle \tag{14}$$

where $\xi_1 \circledast \xi_2 \circledast \ldots \circledast \xi_n$ is the symmetric tensor product [5,6]. Since finite linear combinations of such tensor products are dense in $H_1^{n\circledast}$, it is now easy to see that $H_1^{n\circledast}$ is isometrically isomorphic to $\hat{L}_2(\mathbb{R}^n)$ and hence to H_1^{n*}. For $n=2$ we can illustrate this by factoring the symmetric bilinear map $(\eta_1,\eta_2) \to \eta_1 * \eta_2$ as follows:

$$\begin{tikzcd} H_1 \times H_1 \arrow[rr, "*"] \arrow[dr, "\circledast"'] & & H_2 = H_1 * H_1 \\ & H_1 \circledast H_1 \arrow[ur, "\phi_2"'] & \end{tikzcd}$$

where ϕ_2 is the unique linear map which makes the diagram commute; ϕ_2 is unitary. Similar unitary maps ϕ_n are defined for $n=3,4,\ldots$

If $A_1, A_2, \ldots,$ are linear operators in H_1, we define $A_1 * A_2 * \ldots * A_n : H_n \to H_n$ via

$$(A_1 * A_2 * \ldots * A_n)(\eta_1 * \eta_2 * \ldots * \eta_n) = (A_1\eta_1) * (A_2\eta_2) * \ldots * (A_n\eta_n)$$

on a dense set in H_n and then extend it continuously to all of H_n. We define $A_1 \circledast A_2 \circledast \ldots \circledast A_n : H_1^{n\circledast} \to H_1^{n\circledast}$ analogously. For $n=2$ we have then the following picture:

$$\begin{tikzcd} H_2 \arrow[r, "A_1 * A_2"] \arrow[d, "\phi_2"'] & H_2 \arrow[d, "\phi_2"] \\ H_1 \circledast H_1 \arrow[r, "A_1 \circledast A_2"] & H_1 \circledast H_1 \end{tikzcd}$$

and analogously for $n > 2$.

4. STATE SPACE CONSTRUCTION

THEOREM 1. *The subspace* $X \subset H$ *is a minimal state space for* y *if and only if*

$$X = \bigoplus_{n \in N} X_n \tag{15a}$$

where X_1 *is a minimal Markovian* (H_1^-, H_1^+) *-splitting subspace and*

$$X_n = X_1^{n*} := X_1 * X_1 * \dots * X_1 \tag{15b}$$

(n times). Then each X_n *is a minimal* (H_n^-, H_n^+) *-splitting subspace.*

The proof of this theorem is based on the following lemmas.

LEMMA 1. *Let* $\eta = \eta_1 \circledast \eta_2 \circledast \dots \circledast \eta_n$ *where* $\eta_i \in H_1$ *for* $i=1,2,\dots,n$. *Let* X *be a subspace of* H_1. *Then*

$$E^{X \circledast X \circledast \dots \circledast X}\eta = (E^X\eta_1) \circledast (E^X\eta_2) \circledast \dots \circledast (E^X\eta_2) .$$

PROOF. Let $\hat{\eta}_i := E^X\eta_i$, and let $\xi = \xi_1 \circledast \xi_2 \circledast \dots \circledast \xi_n$ where $\xi_1, \xi_2, \dots, \xi_n$ are arbitrary elements in X. Then, by (14),

$$\langle \eta_1 \circledast \eta_2 \circledast \dots \circledast \eta_n - \hat{\eta}_1 \circledast \hat{\eta}_2 \circledast \dots \circledast \hat{\eta}_n \rangle_{n\circledast} =$$

$$\langle (\eta_1 - \hat{\eta}_1) \circledast \eta_2 \circledast \dots \circledast \eta_n, \xi \rangle_{n\circledast} + \langle \hat{\eta}_1 \circledast (\eta_2 - \hat{\eta}_2) \circledast \dots \circledast \eta_n, \xi \rangle_{n\circledast} + \dots + \langle \hat{\eta}_1 \circledast \hat{\eta}_2 \circledast \dots \circledast (\eta_n - \hat{\eta}_n), \xi \rangle_{n\circledast} =$$

$$\frac{1}{n!} \sum_{\pi \in G} \{ \langle \eta_1 - \hat{\eta}_1, \xi_{\pi_1} \rangle \langle \eta_2, \xi_{\pi_2} \rangle \dots \langle \eta_n \xi_{\pi_n} \rangle + \langle \eta_1, \xi_{\pi_1} \rangle \langle \eta_2 - \hat{\eta}_2, \xi_{\pi_2} \rangle \dots \langle \eta_n, \xi_{\pi_n} \rangle +$$

$$\dots + \langle \eta_1, \xi_{\pi_1} \rangle \langle \eta_2, \xi_{\pi_n} \rangle \dots \langle \eta_n - \hat{\eta}_n, \xi_{\pi_n} \rangle \} ,$$

which equals zero since $\eta_i - \hat{\eta}_i \perp X$. □

LEMMA 2. *Let* X_1 *be a subspace of* H_1, *and let* X_n *be defined by* (15b). *Then* X_1 *is an* (H_1^-, H_1^+)*-splitting subspace if and only if* X_n *is an* (H_n^-, H_n^+)*-splitting subspace.*

PROOF. Due to isomorphism, we can identify H_n^-, H_n^+ and X_n with $(H_1^-)^{n\circledast}$, $(H_1^+)^{n\circledast}$ and $X_1^{n\circledast}$ respectively. But Lemma 1, (14), and the definition of splitting subspace imply that $H_1^- \perp H_1^+ \mid X$ if and only if $(H_1^-)^{n\circledast} \perp (H_1^+)^{n\circledast} \mid X_1^{n\circledast}$. Hence the lemma follows. □

LEMMA 3. *Let* X_1 *and* X_n *be the splitting subspaces of Lemma 2. Then* X_n *is minimal if and only if* X_1 *is minimal.*

PROOF. By Proposition 1 in [12] it suffices to show that the condition $\bar{E}^{X_1}H_1^- = X_1$ is equivalent to $\bar{E}^{X_n}H_n^- = X_n$ (bar over E stands for closure) and that $\bar{E}^{X_1}H_1^+ = X_1$ is equivalent to $\bar{E}^{X_n}H_n^+ = X_n$. By isomorphism $\bar{E}^{X_n}H_n^- = X_n$ can be identified with $\bar{E}^{X_1^{n\circledast}}(H_1^-)^{n\circledast} = X_1^{n\circledast}$, which, by Lemma 1, holds if and only if $\bar{E}^{X_1}H_1^- = X$. A similar argument establishes the other equivalence. □

PROOF OF THEOREM 1. Let the X described in the theorem be denoted $\hat{X}$, and set $\hat{\mathcal{X}} := \sigma(\hat{X}_1)$. Then $\hat{X}$ is the space of all centered $\hat{\mathcal{X}}$-measurable random variables in H [5,6], and

$\sigma(\hat{X}) = \hat{\mathcal{X}}$. Moreover, if $X^- := \bigvee_{t\le 0} U_t X$, it is not hard to see that $\hat{\mathcal{X}}^- := \sigma(\hat{X}^-) = \sigma(\hat{X}_1^-)$, since $U_t \hat{X}_n = (U_t X_1)^{n*}$. An analogous relation holds for summation over the future.

(if): Show that $\hat{X}$ is a minimal state space. Since $\hat{X}_1$ is Markovian, so is $\hat{X}$. Hence, in view of Lemma 2, $\hat{X}$ is a state space for y. Now assume that $\hat{X}$ is not minimal. Then there is a state space X such that $\mathcal{X} := \sigma(X)$ is properly contained in $\hat{\mathcal{X}}$. Then, since all $\xi \in X$ are $\hat{\mathcal{X}}$-measurable and $\hat{X}$ is the space of all centered $\hat{\mathcal{X}}$-measurable random variables, $X \subset \hat{X}$. Therefore X_1 must be a proper subset of $\hat{X}_1$, or else $\mathcal{X} \supset \sigma(X_1) = \sigma(\hat{X}_1) = \hat{\mathcal{X}}$. This contradicts the minimality of $\hat{X}_1$.

(only if): Let X be a minimal state space for y. First let us assume that X_1 is not minimal. Then there is another (H_1^-, H_1^+)-splitting subspace $\hat{X}_1$ which is a proper subspace of X_1. Let $\hat{\mathcal{X}} = \sigma(\hat{X}_1)$, and let $\hat{X}$ be the space of all $\hat{\mathcal{X}}$-measurable elements in H. Clearly $\hat{\mathcal{X}} \subset \mathcal{X} := \sigma(X)$. We want to show that this inclusion is proper, contradicting minimality of X. But this is the case, for there is a $\xi \in X_1$ such that $\xi \perp \hat{X}_1$. Consequently, by the Gaussian property, $\sigma\{\xi\}$ and $\hat{\mathcal{X}}$ are independent, while both are subfields of $\mathcal{X}$. Hence X_1 must be minimal, and $X_1 = \hat{X}_1$. Next assume that X_n is not of the form (15b), i.e. $X_n \neq \hat{X}_n$. Then since $\hat{X}_n$ is minimal (Lemma 3), $X_n \not\subset \hat{X}_n$, i.e. there is a $\xi \in X$ which does not belong to $\hat{X}$ and consequently is not $\hat{\mathcal{X}}$-measurable. Hence $\hat{\mathcal{X}}$ is a proper subfield of $\mathcal{X}$ contradicting minimality of X. Therefore $X = \hat{X}$. Finally $\hat{X}$ is Markovian only if $\hat{X}_1$ is Markovian. The last statement of the theorem follows from Lemma 3. □

5. THE STATE SPACE COMPONENT OF THE FIRST CHAOS

Thus it remains to determine the minimal Markovian (H_1^-, H_1^+)-splitting subspaces X_1. This is *almost* the problem solved in [1-3]. To explain how it differs, let $\zeta \in H_1^- \cap H_1^+$ be defined in the following manner. If $y_1(0) \neq 0$, set $\zeta := y_1(0)$, otherwise let it be arbitrary. (Remember that $H_1^- \cap H_1^+ \neq \emptyset$.) Next define the process $z(t) := U_t \zeta$. Then $z(t) \in H_1$ for all t. Moreover,

$$\begin{cases} H_1^-(z) \subset H_1^- & (16a) \\ H_1^+(z) \subset H_1^+ & (16b) \end{cases}$$

where $H_1^-(z)$ and $H_1^+(z)$ are the closed linear hulls of the random variables $\{z(t); t \le 0\}$ and $\{z(t); t \ge 0\}$ respectively. Since y is purely nondeterministic and mean-square continuous, so is z. Therefore z has a spectral density $\Phi(i\omega)$. A scalar solution W of the equation

$$W(s)W(-s) = \Phi(s) \tag{17}$$

will be called a (full-rank) *spectral factor* of z. Now, if y is Gaussian as assumed in [1,2], $z = y$ and we have equality in each of relations (16). Then there is a procedure in [1,2] to determine X_1 from a certain pair $(W, \bar{W})$ of spectral factors. However, in the non-Gaussian case, $z \neq y$, and we cannot assume that relations (16) hold with equality, not even when $z = y_1$. Hence there is a "mismatch" between the process z and the geometry in H_1, and consequently the procedure of [1,2] will have to be modified.

Fortunately the basic results of [1,2] depend in no crucial way on the spectral factor construction. The following result found in [1,2] is a consequence of the geometry in H_1 only. The theorem requires some new notation: For any Wiener process $\{u(t); t \in \mathbb{R}\} \subset H_1$, let $H_1^-(du)$ and $H_1^+(du)$ be the Gaussian spaces generated by the increments $\{u(\tau)-u(\sigma); \tau,\sigma \le 0\}$ and $\{u(\tau)-u(\sigma); \tau,\sigma \ge 0\}$ respectively. In particular, we have $H_1^-(d\nu) = H_1^-$ and $H_1^+(d\bar{\nu}) = H_1^+$. Here and in the sequel, when we talk of a "Wiener process," we shall always mean a centered Gaussian process *defined on the whole real line* by a spectral representation

$$u(t) = \int_{-\infty}^{\infty} \frac{e^{i\omega t}-1}{i\omega}\, d\hat{u}(i\omega) \ , \tag{18}$$

where $d\hat{u}$ is a Gaussian orthogonal stochastic measure such that $E|d\hat{u}|^2 = \frac{1}{2\pi}\, d\omega$.

THEOREM 2. *A subspace* $X_1 \subset H_1$ *is a minimal Markovian* (H_1^-, H_1^+)*-splitting subspace if and only if*

$$X = H_1^-(du) \ominus H_1^-(d\bar{u}) \tag{19}$$

for some pair $(u,\bar{u})$ *of Wiener processes in* H_1 *such that*

$$H_1^-(d\bar{u}) \subset H_1^-(du) \tag{20a}$$

$$H_1^- \subset H_1^-(du) \tag{20b}$$

$$H_1^+ \subset H_1^+(d\bar{u}) \tag{20c}$$

$$H_1^+(d\bar{u}) = H_1^+ \vee H_1^+(du) \tag{20d}$$

$$H_1^-(du) = H_1^- \vee H_1^-(d\bar{u}) \ . \tag{20e}$$

The processes u and $\bar{u}$ (which are essentially unique) are called respectively the *forward* and the *backward generating processes* of X. (Condition (20a) is equivalent to $H_1^-(du)$ and $H_1^+(d\bar{u})$ intersecting perpendicularly. Moreover, (20d) is an observability and (20e) a constructibility condition [1,2].)

The Gaussian space of any Wiener process u in H_1 coincides with H_1 [9], and consequently any $\eta \in H_1$ can be written

$$\eta = \int_{-\infty}^{\infty} f(-t)\,du(t) \tag{21a}$$

where $f \in L_2(\mathbb{R})$, or equivalently,

$$\eta = \int_{-\infty}^{\infty} \hat{f}(i\omega)\,d\hat{u}(i\omega) \tag{21b}$$

where $\omega \to \hat{f}(i\omega)$ is the Fourier transform [9]. [We shall refer also to the function $\hat{f}$ as the Fourier-transform, although it properly should be called the (double-sided) Laplace-transform.] Relations (22) establishes an isometric isomorphism between H_1 and $L_2(\mathbb{I})$, where $\mathbb{I}$ is the imaginary axis. Let $T_u : H_1 \to L_2(\mathbb{I})$ be the map $T_u\eta = \hat{f}$. Then it can be seen that T_u is unitary. Let T_u^* denote the adjoint, i.e. $\eta = T_u^*\hat{f}$, which is relation (21b). The shift U_t corresponds to $e^{i\omega t}$ under the isomorphism T_u.

LEMMA 4. *There is a one-one correspondence between Wiener processes* u *in* H_1 *and spectral factors* W *of* z *described by the following rule. For each* u, $W := T_u\zeta$ *is a spectral factor. For each spectral factor* W, u *defined by* (18) *and* $d\hat{u} = W^{-1}Gd\hat{v}$, *where* $G := T_v\zeta$, *is a Wiener process.*

PROOF. Let $W := T_u\zeta$. Then

$$z(t) = \int_{-\infty}^{\infty} e^{i\omega t} W d\hat{u} \ , \tag{22}$$

from which it is easy to see that the inverse Fourier transform of $E\{z(t)z(0)'\}$ is $W(i\omega)W(-i\omega)$, establishing W as a spectral factor. In particular G is a spectral factor. Therefore, for any spectral factor W, $d\hat{u} := W^{-1}Gd\hat{v}$ is a Gaussian orthogonal stochastic measure such that $E|d\hat{u}|^2 = \frac{1}{2\pi}d\omega$, for $d\hat{v}$ is. Hence (18) is a Wiener process. □

Next we introduce the *Hardy spaces* H_2^+ and H_2^-: Let $H_2^+(H_2^-)$ be the subspace of $L_2(\mathbb{I})$ of functions whose inverse Fourier-transforms vanish on the negative (positive) real line. From (21) it follows that $T_uH^-(du) = H_2^+$ and $T_uH^+(du) = H_2^-$. A function K which is bounded and analytic in the open left half-plane and has modulus one on the imaginary axis is called *inner*. Define $K^*(i\omega) := K(-i\omega)$; K^* is the inverse of K. If $f \in H_2^+$ and K is inner, $fK \in H_2^+$, and H_2^+K is a subspace of H_2^+. Let $H(K)$ denote the orthogonal complement of H_2^+K in H_2^+.

THEOREM 3. *Let* $\zeta \in H_1^- \cap H_1^+$ *be arbitrary, and set* $G := T_v\zeta$ *and* $\bar{G} := T_{\bar{v}}\zeta$. *Let* $\Gamma := G/\bar{G}$. *Then* X_1 *is a minimal Markovian* (H_1^-, H_1^+)*-splitting subspace if and only if there is a pair of inner functions* $(Q, \bar{Q}^*)$ *such that* $K := \Gamma Q\bar{Q}^*$ *is also inner,* K *and* Q *are coprime,* K *and* $\bar{Q}^*$ *are coprime, and*

$$X_1 = T_u^* H(K) \ , \tag{23}$$

where u *is the Wiener process* (18) *with* $d\hat{u} = Q^*d\hat{v}$.

PROOF. We present an appropriately modified version of the proof in [1,2]. The idea is to translate conditions (20) to the Hardy space setting and apply Beurling's Theorem [13]. To this end, first note that if u_1 and u_2 are two Wiener processes in H_1, and W_1 and W_2 are their corresponding spectral factors,

$$T_{u_2}\eta = (T_{u_1}\eta)(W_2/W_1) \tag{24}$$

for any $\eta \in H_1$, as is easily seen from Lemma 4. Then, if W and $\bar{W}$ are the spectral factors corresponding to u and $\bar{u}$ respectively, applying the map T_u to (20a), (20b) and (20e) and $T_{\bar{u}}$ to (20c) and (20d) yields

$$H_2^+K \subset H_2^+ \quad \text{where} \quad K := W/\bar{W} \tag{25a}$$

$$H_2^+Q \subset H_2^+ \quad \text{where} \quad Q := W/G \tag{25b}$$

$$H_2^-\bar{Q} \subset H_2^- \quad \text{where} \quad \bar{Q} := \bar{W}/\bar{G} \tag{25c}$$

$$H_2^- = (H_2^-\bar{Q}) \vee (H_2^-K^*) \tag{25d}$$

$$H_2^+ = (H_2^+Q) \vee (H_2^+K) \ . \tag{25e}$$

Now, since $H_1^-(d\bar{u})$ is invariant under the shift $\{U_t; t \le 0\}$, H_2^+K is invariant under $\{e^{i\omega t}; t \le 0\}$. Therefore, by Beurling's Theorem [13], (25a) holds if and only if K is inner. In the same way we see that (25b) is equivalent to Q being inner and (25c) to $\bar{Q}$ being inner with respect to H_2^-, i.e. $\bar{Q}^*$ inner. Moreover, (25d) and (25e) are valid if and only if the stated coprimeness conditions hold [13], and (19) is equivalent to $T_u X = H_2^+ \ominus (H_2^+ K) =: H(K)$, i.e. (23). The statement about u follows from Lemma 4. □

REMARK. Let us pinpoint in what way this theorem differs from the corresponding result in [1,2]. In the case studied in [1,2], the pairs $(W,\bar{W})$ which generate splitting subspaces are precisely those for which $W \in H_2^+$, $\bar{W} \in H_2^-$, and K is inner. In the present setting these three conditions must also hold, but in addition we must have $W = GQ$ and $\bar{W} = \bar{G}\bar{Q}$. These factorizations correspond to the inner-outer factorizations of [1,2], but the difference is that now G and $\bar{G}$ are not outer. Consequently, some of the pairs $(W,\bar{W})$ mentioned above will be excluded. Note that the innovation process does not correspond to an outer spectral factor of z, since $\bar{E}^{H^-}H^+$ is not the predictor space of z. □

6. THE STATE PROCESS

We recall from Section 2 that

$$y(t) = \sum_{n \in N} y_n(t) \tag{26a}$$

where

$$y_n(t) = \int_{-\infty}^{t} \int_{-\infty}^{t_1} \cdots \int_{-\infty}^{t_{n-1}} g_n(t-t_1, t-t_2, \ldots, t-t_n)\, d\nu(t_1)\, d\nu(t_2) \ldots d\nu(t_n) \tag{26b}$$

for some $g_n \in L_2(\mathbb{R}^n)$. Let us assume that this innovation representation is given, i.e. that the functions $\{g_n; n \in N\}$ are known.

Let us now consider a minimal state space X with forward generating process u. Then, since $H_1^- \subset H_1^-(du)$,

$$y_n(0) \in H_n^-(du) := H_1^-(du) * H_1^-(du) * \cdots * H_1^-(du) \tag{27}$$

(n times) and consequently there is a representation

$$y_n(0) = \int_{-\infty}^{0} \int_{-\infty}^{t_1} \cdots \int_{-\infty}^{t_{n-1}} w_n(-t_1, -t_2, \ldots, -t_n)\, du(t_1)\, du(t_2) \ldots du(t_n) \tag{28}$$

for some $L_2(\mathbb{R}^n)$. Defining w_n to be zero whenever some argument is zero, we may write this $y_n(0) = I_n(w_n; u)$. By the same recipe we write $y_n(0) = I_n(g_n; \nu)$. We need to determine w_n from g_n. To this end, let $\hat{f} \in L_2(\mathbb{I}^n)$ be the n-fold Fourier-transform

$$\hat{f}(i\omega_1, \ldots, i\omega_n) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-i\omega_1 t_1 - \cdots - i\omega_n t_n} f(t_1, \ldots, t_n)\, dt_1 \ldots dt_n .$$

Let $F_n : L_2(\mathbb{R}^n) \to L_2(\mathbb{I}^n)$ be the operator defined by $\hat{f} = F_n f$. The following is a multidimensional version of (21).

LEMMA 5. *Let* $f \in L_2(\mathbb{R}^n)$ *and set* $\hat{f} := F_n f$. *Let* u *be a Wiener process* (18). *Then*

$$I_n(f; u) = I_n(\hat{f}; \hat{u}) . \tag{29}$$

PROOF. First let f be of the form (8). Then $\hat{f}$ has this form too, and $\hat{f}_i = F_1 f_i$. From (21) we have

$$\eta_i = \int_{-\infty}^{\infty} f_i(-t)\,du(t) = \int_{-\infty}^{\infty} \hat{f}_i(i\omega)\,d\hat{u}(i\omega) \ .$$

Then, by Itô's formula, i.e. (11) with ν exchanged for u or $\hat{u}$, each member of (29) can be reduced to the same expression in $\eta_1, \eta_2, \ldots, \eta_n$. Hence (29) holds for functions of type (8). Then, since finite linear combinations of functions of type (8) are dense in $L_2(\mathbb{R}^n)$ or $L_2(\mathbb{I}^n)$, (20) holds in general. □

Consequently, defining $W_n := F_n w_n$ and $G_n := F_n g_n$, w_n can be determined from g_n via the relation

$$W_n(i\omega_1, \ldots, i\omega_n) = G_n(i\omega_1, \ldots, i\omega_n) Q(i\omega_1) \ldots Q(i\omega_n) \ , \tag{30}$$

for $d\hat{\nu} = Q d\hat{u}$ (Theorem 3).

It is well-known [6], and we have already used this fact in Section 2, that $\eta = I_n(\hat{f};\hat{u})$ defines an isomorphism between H_n and $L_2(\mathbb{I}^n)$. More precisely $T_u^{(n)} : H_n \to L_2(\mathbb{I}^n)$, defined by $T_u^{(n)}\eta = \hat{f}$, is a map with the property that $(n!)^{\frac{1}{2}} T_u^{(n)}$ is unitary. The space of Fourier-transforms of functions (such as w_n) in $L_2(\mathbb{R}^n)$ which vanish whenever an argument is negative, can be identified with $(H_2^+)^{n\circledast}$ so that $T_u^{(n)} H_n^-(du) = (H_2^+)^{n\circledast}$. In the sequel we shall use precisely this realization of the tensor-product Hilbert space $(H_2^+)^{n\circledast}$. Then the tensor product $f_1 \circledast f_2 \circledast \ldots \circledast f_n$ is given by (8). Also, for subspaces $A_1, A_2, \ldots, A_n$ in H_n,

$$T_n^{(n)}\{A_1 * A_2 * \ldots * A_n\} = (T_u A_1) \circledast (T_u A_2) \circledast \ldots \circledast (T_u A_n) \tag{31}$$

so that in particular

$$T_u^{(n)} X_n = H(K) \circledast H(K) \circledast \ldots \circledast H(K) \tag{32}$$

(n times). Then, since $y_n(0) \in X_n$, $W_n \in H(K)^{n\circledast}$.

Following [1,2] we say that X is *regular* if $H(K)$ contains only Fourier-transforms of continuous functions. All X with dim $X_1 < \infty$ are clearly regular. It can be shown [1,2] that if X is regular the functional

$$V\hat{f} = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{f}(i\omega_1, \ldots, i\omega_n)\,d\omega_1 \ldots d\omega_n \tag{33}$$

is bounded on $H(K)^{n\circledast}$. Hence, since $V\hat{f} = f(0)$ where $f = F_n^{-1}\hat{f}$, there is a $B_n \in H(K)^{n\circledast}$ such that

$$f(0) = \langle \hat{f}, B_n \rangle_{H(K)^{n\circledast}} \tag{34}$$

(Riesz Theorem). Next, as in [1,2], we define a strongly continuous semigroup $\{e^{At}; t \geq 0\}$ on $H(K)$ by

$$e^{At}\hat{f} = P^{H(K)} e^{-i\omega t}\hat{f} \tag{35}$$

where $P^{H(K)}$ denotes the orthogonal projection on the subspace $H(K)$. Moreover define the linear bounded operator $C_n : H(K)^{n\circledast} \to \mathbf{R}$ given by

$$C_n\hat{f} = \langle W_n, \hat{f}\rangle_{H(K)^{n\circledast}} . \tag{36}$$

Then the following lemma is a multilinear version of the construction in [14,15] which is being used in [1,2].

LEMMA 6. *The integrand in* (28) *admits the factorization*

$$w_n(t_1,t_2,\ldots,t_n) = C_n\left(e^{At_1}\circledast e^{At_2}\circledast\ldots\circledast e^{At_n}\right)B_n \tag{37}$$

for $t_1,t_2,\ldots,t_n \geq 0$.

PROOF. In view of (34)

$$w_n(t_1,t_2,\ldots,t_n) = \left\langle e^{i\omega_1 t_1+i\omega_2 t_2+\ldots+i\omega_n t_n}W_n, B_n\right\rangle .$$

Since $W_n \in H(K)^{n\circledast}$, we have

$$w_n(t_1,t_2,\ldots,t_n) = \left\langle W_n, P^{H(K)^{n\circledast}} e^{-i\omega_1 t_1-i\omega_2 t_2-\ldots-i\omega_n t_n}B_n\right\rangle$$

which is the required result. □

Consequently

$$y_n(t) = C_n x_n(t) \tag{38}$$

where $x_n(t)$ is the $H(K)^{n\circledast}$-valued process

$$x_n(t) = \int_{-\infty}^{t}\int_{-\infty}^{t_1}\ldots\int_{-\infty}^{t_{n-1}}\left[e^{A(t-t_1)}\circledast\ldots\circledast e^{A(t-t_n)}\right]B_n du(t_1)\ldots du(t_n) . \tag{39}$$

If $H(K)$ is infinite dimensional, $\{x_n(t); t\in\mathbb{R}\}$ is not an ordinary stochastic process but must be defined in a weak sense [16]. Then the *state process* $\{x(t); t\in\mathbb{R}\}$ is defined as the (possibly weakly defined) $\oplus_{n\in N} H(K)^{n\circledast}$-valued process with components x_n; $n\in N$. This terminology is motivated by the following result developed along the lines in [1,2].

PROPOSITION 1. *Let* X *be a regular state space and let* x_n *be given by* (39). *Then*

$$\{\langle\hat{f}, x_n(0)\rangle_{H(K)^{n\circledast}} \mid \hat{f}\in H(K)^{n\circledast}\} = X_n . \tag{40}$$

Moreover, for each $n\in N$,

$$X_n \perp H_n^+(du) . \tag{41}$$

PROOF. Let $\xi\in X_n$ be arbitrary, and let $\hat{f} := T_u^{(n)}\xi$ and $f := F_n^{-1}\hat{f}$. Then, by Lemma 5,

$$\xi = \int_{-\infty}^{0}\int_{-\infty}^{t_1}\ldots\int_{-\infty}^{t_{n-1}} f(t_1,\ldots,t_n)du(t_1)\ldots du(t_n) . \tag{42}$$

By exchanging w_n for f in the proof of Lemma 6, we obtain

$$f(t_1,\ldots,t_n) = \left\langle\hat{f}, \left(e^{-At_1}\circledast\ldots\circledast e^{-At_n}\right)B_n\right\rangle_{H(K)^{n\circledast}} . \tag{43}$$

Then (42) and (43) together yield (40). In view of (19), $X_1 \subset H_1^-(du) \perp H_1^+(du)$, from which (41) follows. □

In partular, for n=1 we can write (38) and (39) in the following suggestive form

$$\begin{cases} dx_1 = Ax_1 dt + B_1 du \\ y_1 = C_1 x_1 \end{cases} \tag{44}$$

The higher-chaos subsystems are nonlinear. In the next section we shall illustrate this with an example.

Note that a backward realization for X generated by $\bar{u}$ is obtained by developing the above analysis in $(H_2^-)^{n\circledast}$ rather than in $(H_2^+)^{n\circledast}$. Whereas the forward property is characterized by (41), the backward one is determined by $X_n \perp H_n^-(d\bar{u})$ for each $n \in \mathbb{N}$.

Finally, in the case that X is not regular, other constructions involving rigged Hilbert spaces are possible [19].

7. THE FINITE-DIMENSIONAL BILINEAR CASE

To illustrate our point let us consider the simplest possible nonlinear problem. Let the process y have the innovation representation

$$y(t) = \int_{-\infty}^{t} g_1(t-\sigma)d\nu(\sigma) + \int_{-\infty}^{t}\int_{-\infty}^{\tau} g_2(t-\tau,t-\sigma)d\nu(\sigma)d\nu(\tau) \tag{45a}$$

and the backward innovation representation

$$y(t) = \int_{t}^{\infty} \bar{g}_1(t-\sigma)d\bar{\nu}(\sigma) + \int_{t}^{\infty}\int_{\tau}^{\infty} \bar{g}_2(t-\tau,t-\sigma)d\bar{\nu}(\sigma)d\bar{\nu}(\tau) \ . \tag{45b}$$

Assume that $G_1 := F_1 g_1$ is a rational function which is not identically zero. Then $\bar{G}_1 := F_1\bar{g}_1$ has the same properties, and $y_1 \neq 0$. Moreover y_1 has a rational spectral density, namely $\Phi(s) := G_1(s)G_1(-s)$.

Now, setting $\Gamma := G_1/\bar{G}_1$, find all pairs $(Q,\bar{Q}^*)$ of inner functions such that $K := \Gamma Q\bar{Q}^*$ is inner and coprime with Q and $\bar{Q}^*$. For each such solution form

$$X_1 = \int_{-\infty}^{\infty} H(K)Q^* d\nu \ . \tag{46}$$

Theorem 3 states that the X_1-spaces obtained in this way are precisely the minimal Markovian (H_1^-,H_1^+)-splitting subspaces. In particular, $Q_1 = 1$ yields $X_1 = E^{H_1^-}H_1^+$, and $\bar{Q}_1 = 1$ yields $X_1 = E^{H_1^+}H_1^-$. Since Γ is rational, it can be shown that K must be rational, and consequently X_1 is finite-dimensional [17]. In fact, all X_1 have the same dimension n [1,2]. By using the procedure described in Section 7 of [18] we can determine an n×n-matrix A_1 and an n×1-matrix B_1 from K and a 1×n-matrix C_1 from $W := GQ$ so that

$$\begin{cases} dx_1 = A_1 x_1 dt + B_1 du \\ y_1 = C_1 x_1 \ , \end{cases} \tag{47}$$

where $sp\{x_1(t),\dots,x_n(t)\} = U_t X_1$, $H_1^+(du) \perp X_1$ and

$$u(t) = \int_{-\infty}^{\infty} \frac{e^{i\omega}-1}{i\omega} Q^*(i\omega)\, d\hat{v} \ . \tag{48}$$

To each X_1 there corresponds a minimal state space, namely

$$X = X_1 \oplus X_2 \ , \tag{49a}$$

where

$$X_2 = X_1 * X_1 \ . \tag{49b}$$

Hence, for each t, the $\frac{1}{2}n(n+1)$ random variables $\{x_1^{ij} := x_1^i(t) * x_1^j(t); j \le i\}$ span $U_t X_2$. (Remember that $x_1^{ij} = x_1^{ji}$.) Let $\{x_2(t); t \in \mathbb{R}\}$ be the $\frac{1}{2}n(n+1)$-dimensional stationary vector process with components $x^{ij}(t)$. Applying Itô's differentiation rule [6] to

$$x_1^{ij}(t) = x_1^i(t)x_1^j(t) - E\{x_1^i(t)x_1^j(t)\}$$

we obtain

$$dx_1^{ij}(t) = \sum_{k=1}^{n} [a_{ik}x_1^{kj}(t) + a_{jk}x_1^{ik}(t)]dt + (b_i x_1^j(t) + b_j x_1^i(t))du$$

where a_{ik} and b_i are the components of A_1 and B_1 respectively. Defining the $\frac{1}{2}n(n+1)\times\frac{1}{2}n(n+1)$-matrix A_2 and the $\frac{1}{2}n(n+1)\times n$-matrix B_2 appropriately, this can be written

$$dx_2 = A_2x_2dt + B_2x_1du \ . \tag{50}$$

Integrating this bilinear equation we get an expression of type

$$x_2(t) = \int_{-\infty}^{t}\int_{-\infty}^{\tau} f(t-\tau, t-\sigma)\, du(\sigma)\, du(\tau) \ ,$$

where f is a vector-valued function. Moreover

$$y_2(t) = \int_{-\infty}^{t}\int_{-\infty}^{\tau} w_2(t-\tau, t-\sigma)\, du(\sigma)\, du(\tau) \ ,$$

where w_2 is obtained from g_2 via formula (30). Now, since $y_2(0) \in X_2$, there are real numbers $\{c_k; k=1,2,\dots,\frac{1}{2}n(n+1)\}$ such that

$$w_2(\tau,\sigma) = \sum_k c_k f_k(\tau,\sigma)$$

and these numbers can be determined by known methods. Let C_2 be the $\frac{1}{2}n(n+1)$-dimensional row vector with components c_k. Then

$$y_2(t) = C_2x_2(t) \ . \tag{51}$$

Since $y = y_1 + y_2$,

$$\begin{cases} dx_1 = A_1x_1dt + B_1du \\ dx_2 = A_2x_2dt + B_2x_1du \\ \ \ \ y = C_1x_1 + C_2x_2 \end{cases} \tag{52}$$

is a realization of y, for $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ is a Markov process. Note that even if y_1 were zero we would need to include x_1 is the state process x, for x_2 by itself is not Markov.

Let $\hat{x}$ be the state process corresponding to $X_1 = E^{H_1^-}H_1^+$ (in the coordinate-system of (52)). It is shown in [1-3] that, for any X_1, $E^{H_1^-}x_1(0) = \hat{x}_1(0)$. Therefore, in view of the definition of x_2 and the fact that $E^{H_2^-} = E^{H_1^-} * E^{H_1^-}$, $E^{H_2^-}x_2(0) = \hat{x}_2(0)$. ($E^{H_1^-}$ and $E^{H_2^-}$ applied to a vector means that the projection is performed componentwise.) Consequently the conditional expectation of $x(t)$ given Y_t^- is

$$E^{Y_t^-}x(t) = \hat{x}(t) \tag{53}$$

for any realization (52). For this reason, remembering that the forward generating process of $E^{H_1^-}H_1^+$ is the innovation ν, we may call the system

$$\begin{cases} d\hat{x}_1 = A_1\hat{x}_1 dt + B_1 d\nu \\ d\hat{x}_2 = A_2\hat{x}_2 dt + B_2\hat{x}_1 d\nu \\ \quad y = C_1\hat{x}_1 + C_2\hat{x}_2 \end{cases} \tag{54}$$

the *steady state non-linear filter* of (52), and we have shown that this filter is invariant over the class (52) of minimal realizations. A similar result can be obtained for backward realizations in terms of $\bar{\nu}$.

8. CONCLUDING REMARKS

The purpose of this paper is to investigate the structural aspects of the nonlinear stochastic realization problem and to clarify basis concepts. This is a first step toward a nonlinear realization theory. Hence we have not concerned ourselves with algorithmic aspects of the problem, and our analysis is based on the availability of an innovation representation, the actual determination of which is a nontrivial problem in itself (see [20]).

The question of state space construction needs to be further studied. It could be argued that condition (4) is too restrictive since there could well be $(\oplus_{n\in N}H_n^-, \oplus_{n\in N}H_n^+)$-splitting subspaces which are not of the form (4), having a nonzero angle with some (or even all) H_n. Hence, if we can do without realizations of individual y_n but only need their sum y, it is possible that we are missing state spaces of smaller size.

Our interest in the nonlinear realization problem emanates from its potential value as a conceptual framework for certain classes of nonlinear filtering problems. This will be the topic of a future study.

REFERENCES

1. A. Lindquist and G. Picci, State space models for Gaussian stochastic processes, *Stochastic Systems: The Mathematics of Filtering and Identification and Applications*, M. Hazewinkel and J.C. Willems, Eds., Reidel Publ. Co., 1981.
2. A. Lindquist and G. Picci, Realization theory for multivariate Gaussian processes (to appear).
3. G. Ruckebusch, Theorie géometrique de la représentation markovienne, Thèse de doctorat d'état, Univ. Paris VI, 1980.
4. J.L. Doob, *Stochastic Processes*, John Wiley, New York, 1953.

5. J. Neven, *Processus Aléatoire Gaussiens*, Presses de L'Université de Montréal, 1968.
6. G. Kallianpur, *Stochastic Filtering Theory*, Springer-Verlag, 1980.
7. D. Ocone, *Topics in Nonlinear Filtering Theory*, Ph.D. thesis, M.I.T., December 1980.
8. S.K. Mitter and D. Ocone, Multiple integral expansions for nonlinear filtering, *Proc. 18th Conf. Decision and Control*, Ft. Lauderdale, Florida, December 1979.
9. Yu. A. Rozanov, *Stationary Random Processes*, Holden-Day, 1967.
10. K. Ito, Multiple Wiener integrals, *J. Math. Soc. Japan* 3, 1951, 157-169.
11. H.P. McKean, *Stochastic Integrals*, Academic Press, 1969.
12. A. Lindquist, G. Picci and G. Ruckebusch, On minimal splitting subspaces and Markovian representations, *Math. Systems Theory* 12, 271-279 (1979).
13. H. Helson, *Lecture Notes on Invariant Subspaces*, Academic Press, 1964.
14. J.S. Baras and R.W. Brockett, H^2 functions and infinite dimensional realization theory, *SIAM J. Control* 13(1975), 221-241.
15. J.S. Baras and P. Dewilde, Invariant subspace methods in linear multivariable-distributed systems and lumped-distributed network synthesis, *Proc. IEEE* 64(1976), 160-178.
16. A.V. Balakrishnan, *Applied Functional Analysis*, Springer-Verlag, 1976.
17. P.A. Fuhrmann, *Linear Systems and Operators in Hilbert Space*, McGraw-Hill, 1981.
18. A. Lindquist and G. Picci, Realization theory for multivariate Gaussian processes II: State space theory revisited and dynamical representations of finite-dimensional state spaces, *Proc. 2nd Intern. Conf. on Information Sciences and Systems*, Patras, Greece, July 1979, Reidel Publ. Co., 108-129.
19. J.W. Helton, Systems with infinite dimensional state space: The Hilbert space approach, *Proc. IEEE* 64 (January 1976), 145-160.
20. H.P. McKean, Wiener's theory of nonlinear noise, *Stochastic Differential Equations*, SIAM-AMS Proceedings, American Math. Society, 1973, 191-209.